

# The quantum de Finetti representation for the Bayesian Quantum Tomography and the Quantum Discord

V. S. Shchesnovich<sup>1</sup> and D. S. Mogilevtsev<sup>1,2</sup>

<sup>1</sup>*Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Santo André, SP, 09210-170 Brazil*

<sup>2</sup>*Institute of Physics, Belarus National Academy of Sciences, F.Skarina Ave. 68, Minsk 220072 Belarus*

We point out that the quantum de Finetti representation, unique for infinitely extendable exchangeable systems, assigns a non-zero Quantum Discord to uncorrelated systems and thus cannot serve as an universal prior distribution in the Bayesian Quantum Tomography. This apparent paradox stems from linearity of the Born rule for the probability assignment in Quantum Mechanics, which results in mixing of one's knowledge about the quantum state and the representative of the state in one density matrix.

PACS numbers: 03.65.Wj, 42.50.Lc

Quantum Mechanics is formulated with the concept of quantum state, a density matrix  $\rho(t)$  in general, which gives the probabilities of outcomes of all possible measurements on the system. The probabilities are defined by the Positive Operator Valued Measure (POVM):  $\Pi_\alpha \geq 0$ ,  $\alpha = 1, \dots, K$ ,  $\sum_\alpha \Pi_\alpha = I$  with the probability assignment being linear in the density matrix:  $p_\alpha = \text{Tr}(\Pi_\alpha \rho)$ , which is the well-known result of Gleason's theorem [1] (see also Ref. [2]). By measuring the quorum of observables one can obtain the set of data sufficient for complete characterization of the density matrix. Feasibility of such a procedure was brilliantly demonstrated in the beginning of 1990-es [3–6] (actually, the firstly suggested quantum homodyne tomography scheme was proved to be such a powerful and efficient tool, that the whole field of quantum states/processes reconstruction was aptly nicknamed as “Quantum Tomography” (QT) [7]).

Due to irreversible character and statistical nature of the quantum measurements, to perform the QT the experimentalist needs *an ensemble of systems in identically prepared states*. Then, a statistical estimation procedure can be devised allowing one to estimate the density matrix parameters (e. g. by using the Maximal Likelihood Estimation [8, 9] or by resorting to the full Bayesian Statistical Inference [10–14]). If the number  $N$  of systems in the ensemble is sufficiently large, the result of estimation is expected to converge to the actual density matrix.

The quantum no-cloning theorem rules out duplicating of an unknown quantum state [15], thus the QT is always a process of updating of information about the state (if necessary, updating also the parameters of the measurement device *at the same time* [16]). Such an updating is the main idea behind the Bayesian Statistical Inference method. The convergence property of the likelihood function for a generic measurement on the ensemble (i.e. the multinomial distribution) to the Dirac delta-function in the limit of infinite number of measurements [17] (see also Ref. [11]) results in agreement between different Bayesian experimentalists. The Bayesian approach in the

QT was pioneered by K. R. W. Jones [10], who obtained an upper bound on the accessible information obtainable from measurement of a pure quantum state, when the latter is represented by an invariant prior measure, and indicated a measurement scheme (the so-called isotropic scheme) which saturates this bound asymptotically. The Bayesian approach to the QT for the system consisting of 1/2-spins was extensively studied in Ref. [11], where pure as well as mixed quantum states of such spin systems were considered.

We note that measurements in the QT are not restricted to separate measurements on individual systems of the ensemble. For a finite  $N$  it is more efficient to measure the entire ensemble as a combined system, see Refs. [18–20]. As single members of the ensemble are concerned, it is shown that asymptotically to the order of  $1/N$  the effectiveness of the individual measurements approaches that of the combined scheme [21]. However, the combined measurement on several systems of the ensemble allows one to uncover also the correlations between the individual ensemble members.

The central question in the Bayesian QT is the way to represent one's incomplete knowledge about an ensemble. Using two premises about the ensemble, namely that of mutual exchangeability of the individual systems and infinite extendability of the ensemble, a unique answer to this question is the quantum de Finetti representation, which follows from the classical de Finetti theorem and Gleason's theorem [12, 22]. It also gives the quantum Bayesian rule for updating the probability distribution [13]. The quantum de Finetti representation has the following form

$$\rho^{(N)} = \int d\mu(\rho) \rho^{\otimes N}, \quad \rho^{\otimes N} \equiv \underbrace{\rho \otimes \rho \otimes \dots \otimes \rho}_N, \quad (1)$$

where  $d\mu(\rho) = d\rho P(\rho)$  and  $P(\rho)$  is the probability density. The Bayesian QT then proceeds as follows. The experimentalist's prior knowledge about the state is reflected in the prior probability density  $P(\rho)$ . The Bayes rule for updating the probability density after measure-

ments on the first  $M$  systems with the data  $\Pi_1, \dots, \Pi_M$  (i.e. the POVM measurement results for one or several POVMs) reads

$$P(\rho|\Pi_1, \dots, \Pi_M) = \frac{P(\Pi_1, \dots, \Pi_M|\rho)P(\rho)}{P(\Pi_1, \dots, \Pi_M)}, \quad (2)$$

where

$$P(\Pi_1, \dots, \Pi_M) = \int d\rho P(\Pi_1, \dots, \Pi_M|\rho)P(\rho). \quad (3)$$

Then, the quantum state for the remaining systems of the ensemble becomes [13]

$$\rho^{(N-M)} = \int d\rho P(\rho|\Pi_1, \dots, \Pi_M) \rho^{\otimes N-M}. \quad (4)$$

Notice that using the quantum de Finetti representation, (1) or (4), one makes predictions also for the correlations between the individual systems of the ensemble. As noted in Ref. [12], the exchangeable representation (1) cannot be carried to the probability theory formulated in the linear space either over the field of real or quaternionic numbers, thus being a unique feature of the complex Hilbert space. It was also argued that an exchangeable de Finetti state is a natural substitute for the “unknown quantum state” in the QT and that the latter has to be banished from it [12].

However, as we show below, the way the quantum de Finetti representation accounts for correlations of the individual systems of the ensemble bears a considerable problem. The problem lies in the fact that the quantum de Finetti prior for a quantum state of  $N$  exchangeable systems almost surely assigns correlations to them, which manifest themselves in nonzero Quantum Discord (QD) [23]. The QD accounts for non-classical correlations between the individual systems of a composite system and can be experimentally measured (see for instance, Ref. [24]). But as we discuss below, if it is known that there are no such correlations [25] and the information on the measured state is limited (e.g. to basic symmetries) one *cannot combine these two features in a quantum de Finetti prior*. This problem is even worse: the posterior, as given by the exchangeable de Finetti representation Eq. (4), will also have a nonzero QD almost surely (see below).

Let us recall the QD definition [23]. The QD quantifies non-classical correlations of two systems  $A$  and  $B$  of a composite system in the Hilbert space  $\mathcal{H}_A \otimes \mathcal{H}_B$ . Given a density matrix  $\rho$  of a composite state, the QD is the difference between two versions of the mutual information,  $D_A(\rho) = I(\rho) - Q_A(\rho)$ . One is  $I(\rho) = H(\rho_A) + H(\rho_B) - H(\rho)$ , where  $H$  is the von Neumann entropy  $H(\rho) = -\text{Tr}(\rho \ln \rho)$  and  $\rho_{A,B} = \text{Tr}_{A,B}(\rho)$  are the reduced density matrices. The other one is defined by optimizing over all possible measurements in  $A$  and is given as follows  $Q_A(\rho) = H(\rho_B) - \min \sum_k p_k H(\rho_{B|k})$ , where  $\rho_{B|k} =$

$\text{Tr}_A(E_k \otimes \mathbb{1}_B \rho) / \text{Tr}(E_k \otimes \mathbb{1}_B \rho)$  is the state of  $B$  conditioned on outcome  $k$  in  $A$ , and  $\{E_k\}$  is the set of POVM elements. These two formulations of the mutual information are two quantum generalizations of the classical mutual information  $I(A : B) = (HA) + H(B) - H(A, B)$ . On the other hand, the state is of zero QD if and only if there exist a von Neumann POVM  $\Pi_k = |\psi_k\rangle\langle\psi_k|$  that

$$\sum_k (\Pi_k \otimes \mathbb{1}_B) \rho (\Pi_k \otimes \mathbb{1}_B) = \rho, \quad (5)$$

i.e. the  $\rho$  is a state obtainable by a von Neumann measurement, where only classical correlations remain. The QD is currently thought of as a resource for various classically intractable tasks including the quantum computation [26–28]. For instance, the remote state preparation (a variant of the quantum teleportation protocol) based on the QD only, i.e. without the quantum entanglement, was already implemented experimentally [29].

Let us now inspect the QD of the exchangeable state (1). To this goal we invoke a sufficient condition [30] for non-zero QD of a bi-partite quantum system, which is the rank of their correlation matrix  $R_{n,m}$ . In our case the said composite consists of two exchangeable systems, i.e.  $\rho^{(2)} = \sum_{n,m} R_{n,m} A_n \otimes B_m$ , where  $A_n$  and  $B_m$  are

bases in the space of Hermitian matrices acting in  $\mathcal{H}_{A,B}$ . Then  $\text{rank}(R) > d_A = d_B$  implies  $D(\rho) > 0$  (note that for an exchangeable state the QD is symmetric with respect to swapping of systems  $A$  and  $B$ ). Consider the two-dimensional systems, where the calculations are simplified with the help of the Bloch vector representation. In this case, the unique measure of Eq. (1) can be cast as  $d\mu(\rho) = d\mu(\vec{n}) = \frac{3}{4\pi} dn_1 dn_2 dn_3 P(\vec{n})$ , where the Bloch vector satisfies  $\vec{n}^2 \leq 1$ . Hence  $\rho(\vec{n}) = \frac{1}{2}(\mathbb{1} + \vec{n}\vec{\sigma})$ , where  $\vec{\sigma} = \{\sigma_1, \sigma_2, \sigma_3\}$  is the vector of Pauli matrices. Then, by Eq. (1), the density matrix  $\rho^{(1)}$  of system  $A$  or  $B$  is

$$\rho^{(1)} = \int d\mu(\vec{n}) \frac{1}{2}(\mathbb{1} + \vec{n}\vec{\sigma}) = \frac{1}{2}(\mathbb{1} + \vec{x}\vec{\sigma}), \quad (6)$$

where we have denoted  $\vec{x} = \int d\mu(\vec{n}) \vec{n}$ . The composite density matrix reads

$$\rho^{(2)} = \frac{1}{4} \left( \mathbb{1} \otimes \mathbb{1} + \vec{x}\vec{\sigma} \otimes \mathbb{1} + \mathbb{1} \otimes \vec{x}\vec{\sigma} + \sum_{i,j} \tau_{i,j} \sigma_i \otimes \sigma_j \right) \quad (7)$$

with the matrix  $\tau$  defined as  $\tau = \int d\mu(\vec{n}) \vec{n} \otimes \vec{n}$  (in the tensor product notation). Now it is a simple observation that the correlation matrix  $R$  of  $\rho^{(2)}$  given by Eq. (7) in the basis  $(\mathbb{1}, \sigma_1, \sigma_2, \sigma_3)$  has the following block-matrix form

$$R = \frac{1}{4} \begin{pmatrix} 1 & \vec{x}^T \\ \vec{x} & \tau \end{pmatrix}, \quad (8)$$

hence if  $\text{rank}(\tau) = 3 > \dim(\mathcal{H}_A) = 2$  then  $D(\rho^{(2)}) > 0$ . But  $\tau$  is full rank for any distribution  $P(\vec{n})$  provided that

it has three-dimensional domain of support in the Bloch ball. The simplest example of this class is the point-mass distribution with  $P(\vec{n}) = \sum_{\alpha} p_{\alpha} \delta(\vec{n} - \vec{e}_{\alpha})$ , where all  $p_{\alpha} > 0$  and  $\vec{e}_1$ ,  $\vec{e}_2$ , and  $\vec{e}_3$  being *any* three linearly independent Bloch vectors.

Next we evaluate the geometric measure of the QD proposed for the two-qubit system in Ref. [30], which gives the distance to the zero QD states in the Bloch vector space. It reads

$$D(\rho^{(2)}) = \frac{1}{4} (||\vec{x}||^2 + ||\tau||^2 - \lambda_{\max}), \quad (9)$$

where  $||\tau||^2 = \text{Tr}(\tau^T \tau)$  (in our case  $\tau^T = \tau$ ) and  $\lambda_{\max}$  is the maximal eigenvalue of the matrix  $\Lambda = \vec{x}\vec{x}^T + \tau\tau^T$ . Simple calculations give

$$D(\rho^{(2)}) = \frac{1}{4} \min_{\vec{m}_1} \left( \int d\mu(\vec{n}_1) \int d\mu(\vec{n}_2) (1 + \vec{n}_1 \vec{n}_2) \right. \\ \left. \times \vec{n}_1 \left[ \sum_{j=2,3} \vec{m}_j \otimes \vec{m}_j \right] \vec{n}_2 \right), \quad (10)$$

where the vectors  $\vec{m}_j$ ,  $j = 1, 2, 3$ , form an orthonormal basis in the Bloch space. Eq. (10) shows that almost surely  $D(\rho^{(2)}) > 0$ , since the matrix in the square brackets in Eq. (10) is manifestly positive for almost all choices of the measure  $d\mu(\vec{n})$ , while the scalar factor preceding it is also positive. The only exception is the case of a measure  $d\mu(\vec{n})$  which has support on the one-dimensional vector space (parallel to some vector  $\vec{m}_1$ ) i.e. when  $P(\vec{n})$  is *a distribution confined to a line in the Bloch ball*. For the same reason, the nonzero QD will almost surely prevail for any finite number of measurements. This is, in fact, a quite general feature of the QD, since it was shown [31] that the states of zero QD belong to a zero measure subset of all states. Thus, one is led to accept a nonzero QD for the *result* of the Bayesian reconstruction for any finite number of measurements, if the de Finetti exchangeable density matrix is admitted as a prior. The Bayesian experimentalist making measurements on one system at a time will not be confused by this, but if a more advanced set-up is to be used with joint measurements on two or more systems at a time, to measure their correlations, the problem is bound to arise due to the way the exchangeable representation assigns such correlations.

The strength of the Bayesian approach lies in selecting *a judicious prior*, reflecting *all* the information available at hand. This is a recurrent theme of the Bayesian Statistical Inference in general [32, 33]. For instance, if one knows somehow that there are no quantum correlations between the individual systems of the ensemble (e.g. they are created one at a time), one would like to reflect this in the prior information. But since one does not know the exact state of the systems in the ensemble, e.g. only some symmetry considerations are known for

the density matrix parameters, one is forced to select a straight line in the Bloch ball for the prior distribution to have zero QD. But which one should be selected? In the current situation the experimentalist has really no clues for choosing it. A limited prior information on the actual state of the system (a typical case of the QT) and zero QD taken jointly *do not allow* one select a measure in representation Eq. (1) for the state of the ensemble which incorporates all the available information. *Thus, the exchangeable representation (1) is not an universal prior that fits all (even the most typical) cases.*

It is to be noted that the described problem has rather deep roots in the linearity of the Quantum Mechanics itself. The field of Quantum Statistical Inference viewed as a variant of the general Parametric Statistical Inference introduces one special feature: the linearity of the probability assignment on the estimated parameter, i.e. the density matrix  $\rho$  describing the quantum state. This unique feature forces one to mix one's incomplete knowledge on the parameter to be estimated and the parameter itself. This is seen already on the single system level. A prior with the density  $P(\rho)$  for the estimated density matrix  $\rho$  invariably leads to a new density matrix  $\rho_{est} = \int d\rho P(\rho) \rho$  by the total probability assignment

$$p(\Pi_{\alpha}) = \int d\rho P(\rho) p(\Pi_{\alpha}|\rho) = \int d\rho P(\rho) \text{Tr}(\Pi_{\alpha} \rho) \\ = \text{Tr}(\Pi_{\alpha} \rho_{est}), \quad (11)$$

where the passage from the first line to the second is provided by linearity of Born's rule (and convexity of the set of density matrices) with the acceptance of the result  $\rho_{est}$  as the "quantum state" by Gleason's theorem (we note in passing that usage of Gleason's result is an important implicit step in the simple and elegant proof of the quantum de Finetti representation in Ref. [12]). From this point of view, the experimentalist in the process of the QT extracts the actual quantum state from such a mixture. Indeed, the Bayesian updating in the QT is not an unlimited process, but has as a limit the maximal possible information obtainable from the ensemble (for instance, the Jones limit [10] for the pure state QT with the unitarily invariant prior). After the maximal possible information is extracted (to the inevitable imperfections of experimental apparatus and restriction to finite number of measurements) no further update is possible and the experimentalists concludes what is the *actual* state of the ensemble.

In conclusion, we have shown that the exchangeable, i.e. the quantum de Finetti, representation almost surely assigns a nonzero Quantum Discord to ensemble of exchangeable systems. If preparation of uncorrelated systems is assumed, the simultaneous requirements of zero Quantum Discord and exchangeability of the prepared ensemble of quantum states do not allow selection of any prior at all within the quantum de Finetti representation.

Furthermore, we point out that it is the linearity of the Born rule for the probability assignments in Quantum Mechanics that leads to such contradictory requirements since, by Gleason's theorem, one mixes one's knowledge about the state and the representative of the state in one density matrix.

This work was supported by by Belarusian agency BR-RFI (D.M.), and also by the FAPESP and CNPq grant agencies of Brazil.

- 
- [1] A. M. Gleason, J. Math. Mech. **6**, 885 (1957).
  - [2] P. Busch, Phys. Rev. Lett. **91**, 120403 (2003).
  - [3] K. Vogel and H. Risken, Phys. Rev. A **40**, 2847 (1989).
  - [4] D. T. Smithey *et al.*, Phys. Rev. Lett. **70**, 1244 (1993); Phys. Rev. A **48**, 3159 (1993).
  - [5] G. M. D'Ariano, C. Macchiavello, and M. G. A. Paris, Phys. Rev. A **50**, 4298 (1994); G. M. D'Ariano, U. Leonhardt, and H. Paul, Phys. Rev. A **52**, R1801 (1995).
  - [6] T. J. Durra, I. A. Walmsley, and S. Mukamel, Phys. Rev. Lett. **74**, 884 (1995).
  - [7] M. G. A. Paris and J. Řeháček (Eds), *Quantum states estimation*, Lect. Notes Phys. vol. 649 (Springer, Berlin Heidelberg, 2004).
  - [8] Z. Hradil, Phys. Rev. A **55**, R1561 (1997); J. Fiurasek, Z. Hradil, Phys. Rev. A **63**, R020101 (2001); J. Řeháček, Z. Hradil, Phys. Rev. Lett. **88**, 130401 (2002) J. Řeháček, Z. Hradil, M. Zawisky, W. Treimer, M. Strobl, Europhys. Lett. **59** 694 (2002); Z. Hradil, J. Řeháček, Fortschr. Phys. **49** 1083 (2001); M. Jezek, J. Fiurasek, Z. Hradil, Phys. Rev. A **68**, 012305 (2003).
  - [9] Z. Hradil, D. Mogilevtsev, and J. Řeháček, Phys. Rev. Lett. **96**, 230401 (2006); D. Mogilevtsev, J. Řeháček and Z. Hradil, Phys. Rev. A **75**, 012112 (2007); Z. Hradil, D. Mogilevtsev, and J. Rehacek, New J. Phys. **10**, 043022 (2008).
  - [10] K. R. W. Jones, Ann. Phys. **207**, 140 (1991); J. Phys. A: Math. Gen. **24**, L1415 (1991); Phys. Rev. A **50**, 3682 (1994).
  - [11] V. Bužek *et al.*, Ann. Phys. **266**, 454 (1998).
  - [12] C. M. Caves, C. A. Fuchs, and R. Schack, J. Math. Phys. **43**, 4537 (2002).
  - [13] R. Schack, T. A. Brun, and C. M. Caves, Phys. Rev. A **64**, 014305 (2001).
  - [14] R. Blume-Kohout, New J. Phys. **12**, 043034 (2010).
  - [15] W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982).
  - [16] D. Mogilevtsev, Phys. Rev. A **82**, 021807(R) (2010); D. Mogilevtsev; J. Rehacek, Z. Hradil, New J. Phys. **14**, 095001 (2012).
  - [17] B. V. Gnedenko, *The Theory of Probability* (English Translation; Mir Publishers, Moscow, 1978), p. 85.
  - [18] A. Peres and W. K. Wootters, Phys. Rev. Lett. **66**, 1119 (1991).
  - [19] S. Massar and S. Popescu, Phys. Rev. Lett. **74**, 1259 (1995).
  - [20] R. Derka, V. Bužek, and A. K. Ekert, Phys. Rev. Lett. **80**, 1571 (1998).
  - [21] E. Bagan, M. Baig, and R. Muñoz-Tapia, Phys. Rev. Lett. **89**, 277904 (2002).
  - [22] R. L. Hudson and G. R. Moody, Z. Wahrscheinlichkeits-theorieverw. Gebiete **33**, 343 (1976).
  - [23] H. Ollivier and W. H. Zurek, Phys. Rev. Lett. **88**, 017901 (2001).
  - [24] R. Blandino, M. G. Genoni, J. Etesse, M. Barbieri, M. G. Paris, P. Grangier, R. Tualle-Brouiri, Phys. Rev. Lett. **109**, 180402 (2012).
  - [25] For example, the systems may be prepared in the same state one by one sequentially.
  - [26] A. Datta, A. Shaji, and C. M. Caves, Phys. Rev. Lett. **100**, 050502 (2008).
  - [27] D. Cavalcanti *et al.*, Phys. Rev. A **83**, 032324 (2011).
  - [28] V. Madhok and A. Datta, Phys. Rev. A **83**, 032323 (2011).
  - [29] B. Dakić *et al.*, Nat. Phys. **8**, 666 (2012).
  - [30] B. Dakić, V. Vedral, and Č. Brukner, Phys. Rev. Lett. **105**, 190502 (2010).
  - [31] A. Ferraro *et al.*, Phys. Rev. A **81**, 052318 (2010).
  - [32] H. Jeffreys *Scientific Inference*, (Cambridge University Press, Cambridge 1973).
  - [33] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge 2003).